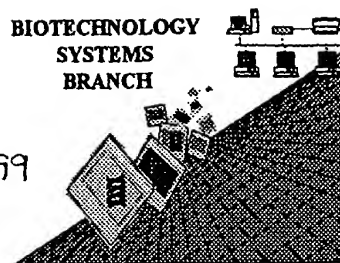


1/a Srivaster -

## NEW SEQUENCE LISTING ERROR REPORT

BIOTECHNOLOGY  
SYSTEMS  
BRANCH



# 8  
5/21/99

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following CRF diskette:

Application Serial Number:

08/962,560

Art Unit / Team No. :

1643

Date Processed by STIC:

2/26/99

**THE ATTACHED PRINTOUT EXPLAINS THE ERRORS DETECTED.**

**PLEASE BE SURE TO FORWARD THIS INFORMATION TO THE APPLICANTS BY EITHER:**

**1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANTS ALONG WITH A NOTICE TO COMPLY or,**

**2) CALLING APPLICANTS AND FAXING THEM A COPY OF THE PRINTOUT WITH A NOTICE TO COMPLY**

**THIS WILL INSURE THAT THE NEXT SUBMISSION RECEIVED FROM THEM WILL BE ERROR FREE.**

**IF YOU HAVE ANY FURTHER QUESTIONS, PLEASE CALL:**

**ARTI SHAH 703-308-4212**

## Raw Sequence Listing Error Summary

### ERROR DETECTED    SUGGESTED    CORRECTION

SERIAL NUMBER:

08/962,560

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1        Wrapped Nucleics      The number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
  
- 2        Wrapped Aminos      The amino acid number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
  
- 3        Incorrect Line Length      The rules require that a line not exceed 72 characters in length. This includes spaces.  
All text must be visible on page.
  
- 4        Misaligned Amino Acid      The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs  
Numbering      between the numbering. It is recommended to delete any tabs and uses spacing between the numbers.
  
- 5        Non-ASCII      This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.  
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
  
- 6        Variable Length      Sequence(s)        contain n's or Xaa's which represented more than one residue.  
As per the rules, each n or Xaa can only represent a single residue.  
Please present the maximum number of each residue having variable length and  
indicate in the (ix) features section that some may be missing.
  
- 7        Wrong Designation      Sequence(s)        contain amino acid or nucleic acid designators which are not standard  
representations as per the Sequence Rules (Please refer to paragraph 1.822)
  
- 8        Skipped Sequences      Sequence(s)        missing. If intentional, please use the following format for each skipped sequence:  
(OLD RULES)      (2) INFORMATION FOR SEQ ID NO:X:  
                                 (i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")  
                                 (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:  
                                 This sequence is intentionally skipped  
  
Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
  
- 9        Skipped Sequences      Sequence(s)        missing. If intentional, please use the following format for each skipped sequence.  
(NEW RULES)      <210> sequence id number  
                                 <400> sequence id number  
                                 000
  
- 10        Use of N's or Xaa's      Use of N's and/or Xaa's have been detected in the Sequence Listing.  
(NEW RULES)      Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
  
- 11        Use of <213>Organism      Sequence(s)        are missing this mandatory field or its response.  
(NEW RULES)
  
- 12        Use of <220>Feature      Sequence(s)        are missing the <220>Feature and associated headings.  
(NEW RULES)      Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"  
                                 (See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32)  
                                 (Sec. 1.823 of new Sequence Rules)
  
- 13        Wrong Format      File submitted was in the alphabetical heading format of the Old Sequence Rules. This is invalid since the  
"Requirements for Patent Applications Containing Nucleotide Sequence and/or Amino Acid Disclosures"  
Federal Register Notice, Vol. 63, No. 104, June 1, 1998, p. 29620  
applies to applications filed on or after July 1, 1998.  
                                 AKS-Biotechnology Systems Branch- 7/10/98

N/A

1643

PAGE: 1

RAW SEQUENCE LISTING  
PATENT APPLICATION US/08/962,560

DATE: 02/26/1999  
TIME: 14:56:37

INPUT SET: S30840.raw

This Raw Listing contains the General  
Information Section and those Sequences  
containing ERRORS.

Does Not Comply  
Corrected Diskette Needed

SEQUENCE LISTING

(1) General Information:

(i) APPLICANT:

*move over - all items 3 on  
Err summary sheet*

Alexandre  
Viney, E  
Willson,  
Richards  
Starr, R  
Nicholson  
Metcalf,  
Nicola,

(ii) TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC AGENTS

(iii) NUMBER OF SEQUENCES: 50

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Scully, Scott, Murphy & Presser  
(B) STREET: 400 Garden City Plaza  
(C) CITY: Garden City, New York  
(D) STATE: New York  
(E) COUNTRY: U.S.A.  
(F) ZIP: 11530

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: ~~US APPLICATION NO.~~ 08/962,560  
(B) FILING DATE: 31-OCT-1997

(vii) ~~+~~ PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/AU97/00729  
(B) FILING DATE: 31-OCT-1997

(viii) ~~+~~ PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: P05117  
(B) FILING DATE: 14-FEB-1997

(vii) PRIOR APPLICATION DATA:

RAW SEQUENCE LISTING  
PATENT APPLICATION US/08/962,560DATE: 02/26/1999  
TIME: 14:56:37

INPUT SET: S30840.raw

46 (A) APPLICATION NUMBER: PO 3384  
47 (B) FILING DATE: 01-NOV-1996  
48  
49  
50 (viii) ATTORNEY/AGENT INFORMATION:  
51 (A) NAME: DiGiglio, Frank S.  
52 (B) REGISTRATION NO: 31,346  
53 (C) REFERENCE/DOCKET NUMBER: 10976  
54  
55 (ix) TELECOMMUNICATION INFORMATION:  
56 (A) TELEPHONE: (516) 742-4343  
57 (B) TELEFAX: (516) 742-4366  
58 (C) TELEX:  
59  
60

## ERRORED SEQUENCES FOLLOW:

61 (2) INFORMATION FOR SEQ ID NO:1:  
62  
63 (i) SEQUENCE CHARACTERISTICS:  
--> 64 (A) LENGTH: 20 base pairs  
65 (B) TYPE: nucleic acid  
66 (C) STRANDEDNESS: single  
67 (D) TOPOLOGY: linear  
68  
69 (ii) MOLECULE TYPE: DNA  
70  
71 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:  
72  
73  
74 CACGCCGCC ACGTGAAGGC  
75  
76

77 (2) INFORMATION FOR SEQ ID NO:2:  
78  
79 (i) SEQUENCE CHARACTERISTICS:  
--> 80 (A) LENGTH: 20 base pairs  
81 (B) TYPE: nucleic acid  
82 (C) STRANDEDNESS: single  
83 (D) TOPOLOGY: linear  
84  
85 (ii) MOLECULE TYPE: DNA  
86  
87 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:  
88  
89 TTCGCCAATG ACAAACGCT  
90  
91

20 ← per sequence  
Rules, insert  
cumulative base  
total at right  
margin of each  
line

same error

INPUT SET: S30840.raw

```

92 (2) INFORMATION FOR SEQ ID NO:3:
93
94 (i) SEQUENCE CHARACTERISTICS:
--> 95 (A) LENGTH: 1236 base pairs
96 (B) TYPE: nucleic acid
97 (C) STRANDEDNESS: single
98 (D) TOPOLOGY: linear
99
100 (ii) MOLECULE TYPE: DNA
101
102 (ix) FEATURE:
103 (A) NAME/KEY: CDS
104 (B) LOCATION: 1..636
105
106 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
107
108
109
--> 110 CGAGGCTCAA GCTCCGGGCG GATTCGCGT GCCGCTCTCG CTCCTGGGG TCTGTGGCC
--> 111 GGCCTGTGCC ACCCGGACGC CCGGCTCACT GCCTCTGTCT CCCCATCAG CGCAGCCCCG
112
--> 113 GACGCTATGG CCCACCCCTC CAGCTGGCCG CTCGAGTAGG
114
115
--> 116 ATG GTA GCA CGC AAC CAG GTG GCA GCC GAC AAT GCG ATC TCC CCG GCA
117 Met Val Ala Arg Asn Gln Val Ala Ala Asp Asn Ala Ile Ser Pro Ala
118 1 5 10 15
119
120 GCA GAG CCC CGA CGG CGG TCA GAG CCC TCC TCG TCC TCG TCT TCG TCC
121 Ala Glu Pro Arg Arg Arg Ser Glu Pro Ser Ser Ser Ser Ser Ser
122 20 25 30
123
124 TCG CCA GCG GCC CCC GTG CGT CCC CGG CCC TGC CCG GCG GTC CCA GCC
125 Ser Pro Ala Ala Pro Val Arg Pro Arg Pro Cys Pro Ala Val Pro Ala
126 35 40 45
127
128 CCA GCC CCT GGC GAC ACT CAC TTC CGC ACC TTC CGC TCC CAC TCC GAT
129 Pro Ala Pro Gly Asp Thr His Phe Arg Thr Phe Arg Ser His Ser Asp
130 50 55 60
131
132 TAC CGG CGC ATC ACG CGG ACC AGC GCG CTC CTG GAC GCC TGC GGC TTC
133 Tyr Arg Arg Ile Thr Arg Thr Ser Ala Leu Leu Asp Ala Cys Gly Phe
134 65 70 75 80
135
136 TAT TGG GGA CCC CTG AGC GTG CAC GGG GCG CAC GAG CGG CTG CGT GCC
137 Tyr Trp Gly Pro Leu Ser Val His Gly Ala His Glu Arg Leu Arg Ala
138 85 90 95
139
140 GAG CCC GTG GGC ACC TTC TTG GTG CGC GAC AGT CGT CAA CGG AAC TGC
141 Glu Pro Val Gly Thr Phe Leu Val Arg Asp Ser Arg Gln Arg Asn Cys
142 100 105 110
143

```

Per 1.822(l) of Sequence Rules

The enumeration of nucleotide  
base shall start at the first base  
of the sequence with number 1.do not use  
negative numbers  
as base totals101 60  
-41 120

-1 160

48 208

↓  
re-number  
ending  
base  
totals

96

144

192

240

288

336

# RAW SEQUENCE LISTING PATENT APPLICATION US/08/962,560

DATE: 02/26/1999  
TIME: 14:56:38

INPUT SET: S30840.raw

144	TTC TTC GCG CTC AGC GTG AAG ATG GCT TCG GGC CCC ACG AGC ATC CGC	384
145	Phe Phe Ala Leu Ser Val Lys Met Ala Ser Gly Pro Thr Ser Ile Arg	
146	115 120 125	
147		
148	GTG CAC TTC CAG GCC GGC CGC TTC CAC TTG GAC GGC AGC CGC GAG ACC	432
149	Val His Phe Gln Ala Gly Arg Phe His Leu Asp Gly Ser Arg Glu Thr	
150	130 135 140	
151		
152	TTC GAC TGC CTT TTC GAG CTG CTG GAG CAC TAC GTG GCG GCG CCG CGC	480
153	Phe Asp Cys Leu Phe Glu Leu Leu Glu His Tyr Val Ala Ala Pro Arg	
154	145 150 155 160	
155		
156	CGC ATG TTG GGG GCC CCG CTG CGC CAG CGC CGC GTG CCG CCG CTG CAG	528
157	Arg Met Leu Gly Ala Pro Leu Arg Gln Arg Val Arg Pro Leu Gln	
158	165 170 175	
159		
160	GAG CTG TGT CGC CAG CGC ATC GTG GCC GCC GTG GGT CGC GAG AAC CTG	576
161	Glu Leu Cys Arg Gln Arg Ile Val Ala Ala Val Gly Arg Glu Asn Leu	
162	180 185 190	
163	GCG CGC ATC CCT CTT AAC CCG GTA CTC CGT GAC TAC CTG AGT TCC TTC	624
164	Ala Arg Ile Pro Leu Asn Pro Val Leu Arg Asp Tyr Leu Ser, Ser Phe	
165	195 200 205	
166		
167	CCC TTC CAG ATC TGA CCGGCTG CCGCTGTGCC GCAGCATTAA GTGGGGGCGC	676
168	Pro Phe Gln Ile *	
169	210	
170		
171	CTTATTATTT CTTATTATTA ATTATTATTA TTTTCTGGA ACCACGTGGG AGCCCTCCCC	736
172		
173	GCCTGGGTCG GAGGGAGTGG TTGTGGAGGG TGAGATGCCT CCCACTTCTG GCTGGAGACC	796
174		
175	TCATCCACC TCTCAGGGGT GGGGGTGCTC CCCTCCTGGT GCTCCCTCCG GGTCCCCCT	856
176		
177	GGTGTAGCA GCTTGTGTCT GGGGCCAGGA CCTGAATTCC ACTCCTACCT CTCCATGTTT	916
178		
179	ACATATTCCC AGTATCTTTG CACAAACCAG GGGTCGGGGA GGGTCTCTGG CTTCAATTTT	976
180		
181	CTGCTGTGCA GAATATCCTA TTTTATATTT TTACAGCCAG TTTAGGTAAT AAACCTTTATT	1036
182		
183	ATGAAAGTTT TTTTAAAAA GAAAAA AAAA	1075
184		
185		
186		

566 (2) INFORMATION FOR SEQ ID NO:9:

567

568 (i) SEQUENCE CHARACTERISTICS:

569 (A) LENGTH: 1094 base pairs

570 (B) TYPE: nucleic acid

571 (D) TOPOLOGY: linear

572

573 (ii) MOLECULE TYPE: ~~protein~~

574

*delete - enabled  
response for a nuclear  
acid sequence*

*(C) STRANDEDNESS: insert  
this  
mandatory  
subheading and  
its response*

RAW SEQUENCE LISTING  
PATENT APPLICATION US/08/962,560DATE: 02/26/1999  
TIME: 14:56:39

INPUT SET: S30840.raw

--> 575 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:  
576  
577 CTCCGGCTGG CCCCTTCTGT AGGATGGTAG CACACAACCA GGTGGCAGCC GACAATGCAG 60  
578 TCTCCACAGC AGCAGAGCCC CGACGGCGGC CAGAACCTTC CTCCTCTTCC TCCTCCTCGC 120  
579 CCGCGGCCCC CGCGCGCCCC CGGCCGTGCC CCGCGGTCCC GGCCCCGGCC CCCGGCGACA 180  
580 CGCACTTCCG CACATTCCGT TCGCACGCCG ATTACCGGCG CATCACGCGC GCCAGCGCGC 240  
581 TCCTGGACGC CTGCGGATTG TACTGGGGGC CCTGAGCGT GCACGGGGCG CACGAGCGGC 300  
582 TGCGCGCCGA GCCCCGTGGG ACCTTCTCTG TGCGCGACAG CCGCCAGCGG AACTGCTTTT 360  
583 TCGCCCTTAG CGTGAAGATG GCCTCGGGAC CCACGAGCAT CCGCGTGCAC TTTCAGGCCG 420  
584 GCCGCTTTCA CCTGGATGGC AGCCGCGAGA GCTTCGACTG CCTCTTCGAG CTGCTGGAGC 480  
585 ACTACGTGGC GCGCGCCGCG CGCATGTCTG GGGCCCCGCT GCGCCAGCGC CGCGTGCGGC 540  
586 CGCTGCAGGA GCTGTGCCGC CAGCGCATCG TGGCCACCGT GGGCCGCGAG AACCTGGCTC 600  
587 GCATCCCCCT CAACCCCGTC CTCCGCGACT ACCTGAGCTC CTTCCCCCTC CAGATTTGAC 660  
588 CGGCAGCGCC CGCCGTGCAC GCAGCATTAA CTGGGATGCC GTGTTATTTT GTTATTACTT 720  
589 GCCTGGAACC ATGTGGGTAC CCTCCCCGGC CTGGGTTGGA GGGAGCGGAT GGGTGTAGGG 780  
590 GCGAGGCGCC TCCCGCCCTC GGCTGGAGAC GAGGCCGCGC ACCCCTTCTC ACCTCTTGAG 840  
591 GGGGTCTCTC CCCTCCTGGT GCTCCCTCTG GGTCCCCCTG GTTGTGTAG CAGCTTAACT 900  
592 GTATCTGGAG CCAGGACCTG AACTCGCACC TCCTACCTCT TCATGTTTAC ATATACCAG 960  
593 TATCTTTGCA CAAACCAGGG GTTGGGGGAG GGTCTCTGGC TTTATTTTTC TGCTGTGCAG 1020  
594 AATCCTATTT TATATTTTTT AAAGTCAGTT TAGGTAATAA ACTTTATTAT GAAAGTTTTT 1080  
595 TTTTTTAAAA AAAA 1094  
596

652 (2) INFORMATION FOR SEQ ID NO:11:

653

654 (i) SEQUENCE CHARACTERISTICS:

655 (A) LENGTH: 2807 base pairs

656 (B) TYPE: nucleic acid

657 (D) TOPOLOGY: linear

658

659 (ii) MOLECULE TYPE: ~~protein~~

660

--&gt; 661 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

662

663

664 GGAAACCGAG GCGGGGAGAC CAGGAGGCCT TGGCCTCAGA GCTTCAGAGT CGCGTGGCAG 60  
665 CAAACAGAGA AACCTGTAGA GGGCAGTGTG CGTCACTTAG CTCAGGGAAG CTGCACGCGA 120  
666 AACTCACCCG CCTTCATTCA TAAACATCGT CAGCTAGGCA CCTACTCCTG GGCTTTCAGG 180  
667 ACAAAGTGAA TCACGAAACC ACAGTGTCCT TAAATAGGT CTGACCGCCT GAATCCCTGG 240  
668 CCAAGGTGTG TACGGGGCAT GGGAGCCCTT GTGCAGAGAT GCTTGCAAGG GCCTTGAGGG 300  
669 GCTCTGTAAG ACAGAGGCTA GGAAGACAAA GTTGGGGGCT ACAGCTTCTT GTCCTGCCCC 360  
670 GGGCCTCAGT TTCTTCGGTT GCCACGTAG GAGTGCAGAG AGTCCAGCCC CTGGGGACCC 420  
671 AACCCTAACCC CGCCAGTTT CCGAGGAAGT CGTCCGGGAG CGGGGGCGCC CCTCCCGCAC 480  
672 CGCCTTAGGC TTCTTTTGAA GCCTCTGCGG TCAGGCCACC GCTTCTCTGGG AAGCCCAAGC 540  
673 CAAGGCCAGG CCGAGTGGCC AACGGGAGGG GCCCGCGCGC GATTCTGGAG GAGGCGGCGC 600  
674 GCCCCACAGG TCTCCAGGGC TGCTAGCCG GGCTCCTAGA GCGGAGACTG CCAAGGCCTT 660  
675 CGGGTCCTGG GCAGGAAGGA TCCTGGCAGG GAGGAGTTGC TTGGGGGGTG GGGGGGAAAG 720  
676 GCTCCAGGCG CGGTGGAGCT CTGACCAGGA GAATGCACAC ACTCGGAGGG GAGGAGGCGT 780  
677 GTCAGCCCCA AGCTAGCATC CCACCCGGGG AGCAGCGATG TGGGGCGAAG GTAGCCAGAG 840  
678 CAAAAGAGCA GGCACCAGGT GACACGAAAC AGAAGATTCC GGGTAGAGCC AGAACCCAG 900  
679 AAGTCCCAT T CAGGAAGGT GCGAGGCGAG AACGAGTTAG GTGGACCTC TCCAGGGGCA 960  
680 GCCAAGAAA TCTAAAGAGA ACCCGAAGGA CTTGCCCGAA AGAGAAACCG AAAGCGGCGG 1020  
681 TGGGCGGGAT CGGTGGGCGG GGCCTCCCTG GTTTAAGAGC TTGATGCAGG GCGGGGCGAG 1080

RAW SEQUENCE LISTING  
PATENT APPLICATION US/08/962,560DATE: 02/26/1999  
TIME: 14:56:39

INPUT SET: S30840.raw

682 AGCAGAGAGA ACTGCGGCCG TGGCAGCGGC ACGGCTCCCC GCCCCGGAGC ATGCGCGACA 1140  
683 GCAGCCCCGG AACCCCCAGC CGCGGCGCCC CGCGTCCCCG CGCCAGGTGA GCCGAGGCAG 1200  
684 CTGCGAAGGA GCAGGCGGGA GGGGATGGGA GGAAGGGGAG CAGAGCCTGG CAGGACTATC 1260  
685 CTCGCAGACT GCATGGCGGG GTCTGTGGATG CTATGCCTCT GGCGCCCCGC CCACCGGCTG 1320  
686 GCCCAGGCGG CCCCTCGCGC GCGCGGGGCG CCGTCAGCCC CTCCTCTCCG GCCCTGAGCC 1380  
687 CGGATCGTCC GCCCGGGTTC CAGTTCCCGG CGTGGCCAGT AGGCGGCAAC CGCGAGGCGG 1440  
688 CAAGCCACCC AGCGGGGACG GCCTGGAGTC GGGCCCCCTCT CCACGCCCCC TTCTCCACGC 1500  
689 GCGCGGGGAG GCAGGGCTCC ACCGCCAGTC TGGAAGGGTT CCACATACAG GAACGGCCTA 1560  
690 CTTGCGAGAT GAGCCCACCG AGGCTCAGGC TCCGGGCGGA TTCTGCGTGT CACCCTCGCT 1620  
691 CTTGGGGTC CGCTGGCCCG CCTGTGCCAC CCGGACGCCC GGTTCACTGC CTCTGTCTCC 1680  
692 CCCATCAGCG CAGCCCCGGA CGCTATGGCC CACCCCTCCA GCTGGCCCTT CGAGTAGGAT 1740  
693 GGTAGCACGT AACCAGGTGG AAGCCGACAA TCGATCTCC CCGGCATCAG AGCCCCGACG 1800  
694 GCGGCCAGAG CCATCCTCGT CCTCGTCTTC GTCCTCGCCG GCGGCCCCCG CGCGTCCCCG 1860  
695 GCCCTGCCCG GTGGTCCCGG CCCC GGCTCC CAGCGACACT CACTTCCGCA CCTTCCGCTC 1920  
696 CCACTCTGAT TACCGGCGCA TCACGCGGAC CAGCGCTCTC CTGGACGCGT GCGGCTTCTA 1980  
697 CTGGGGACCC CTGAGCGTGC ATGGGGCGCA CGAACGGCTG CGTTCCGAAC CCGTGGGCAC 2040  
698 CTTCTTGTTG CGCGACAGTC GCCAGCGGAA CTGCTTCTTC GCGCTCAGCG TGAAGATGGC 2100  
699 TTCGGGCCCC ACGAGCATTC GTGTGCACTT CCAGGCGGCG CGCTTCCACC TGGACGGCAA 2160  
700 CCGCGAGACC TTCGACTGCC TCTTCGAGCT GCTGGAGCAC TACGTGGCGG CGCCGCGCCG 2220  
701 CATGTTGGGG GCCCCACTGC GCCAGCGCCG CGTGGCGCCG CTGCAGGAGC TGTGTGCCCA 2280  
702 GCGCATCGTG GCCGCCGTGG GTCGCGAGAA CCTGGCACGC ATCCCTCTTA ACCCGGTACT 2340  
703 CCGTGACTAC CTGAGTTCCT TCCCCTTCCA GATCTGACCG GCTGCCGCGG TGCCCGCAGA 2400  
704 ATTAAGTGGG AGCGCCTTAT TATTTCTTAT TATTAATTAT TATTATTTT CTGGAACCAC 2460  
705 GTGGGAGCCC TCCCGCGCTA GGTCCGAGGG AGTGGGTGTG GAGGGTGAGA TCCCTCCAC 2520  
706 TTCTGGCTGG AGACCTTATC CCGCTCTCG GGGGGCCTCC CCTCCTGGTG CTCCCTCCCG 2580  
707 GTCCCCCTGG TTGTAGCAGC TTGTGTCTGG GGCCAGGACC TGAACCTCAC GCCTACCTCT 2640  
708 CCATGTTTAC ATGTTCCAG TATCTTTGCA CAAACCAGGG GTGGGGGAGG GTCTCTGGCT 2700  
709 TCATTTTCT GCTGTGCAGA ATATTCTATT TTATATTTT ACATCCAGTT TAGATAATAA 2760  
710 ACTTTATTAT GAAAGTTTTT TTTTTTAAAG AAACAAAGAT TTCTAGA 2807  
711

904 (2) INFORMATION FOR SEQ ID NO:14:

905

906 (i) SEQUENCE CHARACTERISTICS:

--&gt;

907 (A) LENGTH: 420 amino acids

908 (B) TYPE: amino acid

909 (D) TOPOLOGY: linear

910 (ii) MOLECULE TYPE: protein

911

912 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

913

914 Met Ala Ser Phe Pro Pro Arg Val Asn Glu Lys Glu Ile Val Arg Ser

915 1 5 10 15

916 Arg Thr Ile Gly Glu Leu Leu Ala Pro Ala Ala Pro Phe Asp Lys Lys

917 20 25 30

918 Cys Gly Gly Glu Asn Trp Thr Val Ala Phe Ala Pro Asp Gly Ser Tyr

919 35 40 45

920 Phe Ala Trp Ser Gln Gly Tyr Arg Ile Val Lys Leu Val Pro Trp Ser

921 50 55 60

922

923

924

925

*see last page*



RAW SEQUENCE LISTING  
PATENT APPLICATION US/08/962,560DATE: 02/26/1999  
TIME: 14:56:39

INPUT SET: S30840.raw

926	Gln	Cys	Arg	Lys	Asn	Phe	Leu	Leu	His	Gly	Ser	Lys	Asn	Val	Thr	Asn
927	65					70					75					80
928																
929	Ser	Ser	Cys	Leu	Lys	Leu	Ala	Arg	Gln	Asn	Ser	Asn	Gly	Gly	Gln	Lys
930				85						90					95	
931																
932	Asn	Lys	Pro	Pro	Glu	His	Val	Ile	Asp	Cys	Gly	Asp	Ile	Val	Trp	Ser
933				100					105					110		
934																
935	Leu	Ala	Phe	Gly	Ser	Ser	Val	Pro	Glu	Lys	Gln	Ser	Arg	Cys	Val	Asn
936			115					120					125			
937																
938	Ile	Glu	Trp	His	Arg	Phe	Arg	Phe	Gly	Gln	Asp	Gln	Leu	Leu	Leu	Ala
939	130						135					140				
940																
941	Thr	Gly	Leu	Asn	Asn	Gly	Arg	Ile	Lys	Ile	Trp	Asp	Val	Tyr	Thr	Gly
942	145					150					155					160
943																
944	Lys	Leu	Leu	Leu	Asn	Leu	Val	Asp	His	Ile	Glu	Met	Val	Arg	Asp	Leu
945					165					170					175	
946																
947	Thr	Phe	Ala	Pro	Asp	Gly	Ser	Leu	Leu	Leu	Val	Ser	Ala	Ser	Arg	Asp
948				180					185					190		
949																
950	Lys	Thr	Leu	Arg	Val	Trp	Asp	Leu	Lys	Asp	Asp	Gly	Asn	Met	Val	Lys
951			195					200					205			
952																
953	Val	Leu	Arg	Ala	His	Gln	Asn	Trp	Val	Tyr	Ser	Cys	Ala	Phe	Ser	Pro
954		210					215					220				
955																
956	Asp	Cys	Ser	Met	Leu	Cys	Ser	Val	Gly	Ala	Ser	Lys	Ala	Val	Phe	Leu
957	225					230					235					240
958																
959	Trp	Asn	Met	Asp	Lys	Tyr	Thr	Met	Ile	Arg	Lys	Leu	Glu	Gly	His	His
960					245					250					255	
961																
962	His	Asp	Val	Val	Ala	Cys	Asp	Phe	Ser	Pro	Asp	Gly	Ala	Leu	Leu	Ala
963				260					265					270		
964																
965	Thr	Ala	Ser	Tyr	Asp	Thr	Arg	Val	Tyr	Val	Trp	Asp	Pro	His	Asn	Gly
966			275					280					285			
967																
968	Asp	Leu	Leu	Met	Glu	Phe	Gly	His	Leu	Phe	Pro	Ser	Pro	Thr	Pro	Ile
969		290					295					300				
970																
971	Phe	Ala	Gly	Gly	Ala	Asn	Asp	Arg	Trp	Val	Arg	Ala	Val	Ser	Phe	Ser
972	305					310					315					320
973																
974	His	Asp	Gly	Leu	His	Val	Ala	Ser	Leu	Ala	Asp	Asp	Lys	Met	Val	Arg
975					325					330					335	
976																
977	Phe	Trp	Arg	Ile	Asp	Glu	Asp	Cys	Pro	Val	Gln	Val	Ala	Pro	Leu	Ser
978					340				345					350		

# RAW SEQUENCE LISTING PATENT APPLICATION US/08/962,560

DATE: 02/26/1999  
TIME: 14:56:40

INPUT SET: S30840.raw

```

979
980 Asn Gly Leu Cys Cys Ala Phe Ser Thr Asp Gly Ser Val Leu Ala Ala
981      355                      360                      365
982
983 Gly Thr His Asp Gly Ser Val Tyr Phe Trp Ala Thr Pro Arg Gln Val
984      370                      375                      380
985
986 Pro Ser Leu Gln His Ile Cys Arg Met Ser Ile Arg Arg Val Met Ser
987      385                      390                      395                      400
988
989 Thr Gln Glu Val Gln Lys Leu Pro Val Pro Ser Lys Ile Leu Ala Phe
990      405                      410                      415
991
992 Leu Ser Tyr Arg Gly
993      420
994

```

2397 (2) INFORMATION FOR SEQ ID NO:35:

2398

2399 (i) SEQUENCE CHARACTERISTICS:

--> 2400 (A) LENGTH: 848 base pairs  
 2401 (B) TYPE: nucleic acid  
 2402 (C) STRANDEDNESS: single  
 2403 (D) TOPOLOGY: linear

2404

2405 (ii) MOLECULE TYPE: DNA

2406

2407

2408 (ix) FEATURE:

2409 (A) NAME/KEY: CDS  
 2410 (B) LOCATION: 1..624

2411

2412

2413 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

2414

2415 TTG GAG AAG TGT GGT TGG TAT TGG GGG CCA ATG AAT TGG GAA GAT GCA 48

2416 Leu Glu Lys Cys Gly Trp Tyr Trp Gly Pro Met Asn Trp Glu Asp Ala

2417 1 5 10 15

2418

2419 GAG ATG AAG CTG AAA GGG AAA CCA GAT GGT TCT TTC CTG GTA CGA GAC 96

2420 Glu Met Lys Leu Lys Gly Lys Pro Asp Gly Ser Phe Leu Val Arg Asp

2421 20 25 30

2422

2423 AGT TCT GAT CCT CGT TAC ATC CTG AGC CTC AGT TTC CGA TCA CAG GGT 144

2424 Ser Ser Asp Pro Arg Tyr Ile Leu Ser Leu Ser Phe Arg Ser Gln Gly

2425 35 40 45

2426

2427 ATC ACC CAC CAC ACT AGA ATG GAG CAC TAC AGA GGA ACC TTC AGC CTG 192

2428 Ile Thr His His Thr Arg Met Glu His Tyr Arg Gly Thr Phe Ser Leu

2429 50 55 60

2430

2431 TGG TGT CAT CCC AAG TTT GAG GAC CGC TGT CAA TCT GTT GTA GAG TTT 240

2432 Trp Cys His Pro Lys Phe Glu Asp Arg Cys Gln Ser Val Val Glu Phe

*next page*

# RAW SEQUENCE LISTING PATENT APPLICATION US/08/962,560

DATE: 02/26/1999  
TIME: 14:56:40

INPUT SET: S30840.raw

```

2433      65              70              75              80
2434  ATT AAG AGA GCC ATT ATG CAC TCC AAG AAT GGA AAG TTT CTC TAT TTC      288
2435  Ile Lys Arg Ala Ile Met His Ser Lys Asn Gly Lys Phe Leu Tyr Phe
2436              85              90              95
2437
2438  TTA AGA TCC AGG GTT CCA GGA CTG CCA CCA ACT CCT GTC CAG CTG CTC      336
2439  Leu Arg Ser Arg Val Pro Gly Leu Pro Pro Thr Pro Val Gln Leu Leu
2440              100              105              110
2441
2442  TAT CCA GTG TCC CGA TTC AGC AAT GTC AAA TCC CTC CAG CAC CTT TGC      384
2443  Tyr Pro Val Ser Arg Phe Ser Asn Val Lys Ser Leu Gln His Leu Cys
2444              115              120              125
2445
2446  AGA TTC CGG ATA CGA CAG CTC GTC AGG ATA GAT CAC ATC CCA GAT CTC      432
2447  Arg Phe Arg Ile Arg Gln Leu Val Arg Ile Asp His Ile Pro Asp Leu
2448              130              135              140
2449
2450  CCA CTG CCT AAA CCT CTG ATC TCT TAT ATC CGA AAG TTC TAC TAC TAT      480
2451  Pro Leu Pro Lys Pro Leu Ile Ser Tyr Ile Arg Lys Phe Tyr Tyr Tyr
2452  145              150              155              160
2453
2454  GAT CCT CAG GAA GAG GTA TAC CTG TCT CTA AAG GAA GCG CAG CGT CAG      528
2455  Asp Pro Gln Glu Glu Val Tyr Leu Ser Leu Lys Glu Ala Gln Arg Gln
2456              165              170              175
2457
2458  TTT CCA AAC AGA AGC AAG AGG TGG AAC CCT CCA CGT AGC GAG GGG CTC      576
2459  Phe Pro Asn Arg Ser Lys Arg Trp Asn Pro Pro Arg Ser Glu Gly Leu
2460              180              185              190
2461
2462  CCT GCT GGT CAC CAC CAA GGG CAT TTG GTT GCC AAG CTC CAG CTT TGAAGAACCA 631
2463  631
2464  Pro Ala Gly His His Gln Gly His Leu Val Ala Lys Leu Gln Leu
2465              195              200              205
2466
--> 2467  AATTAAGCTA CCATGAAAAG AAGAGGAAAA GTGAGGGAAC AGGAAGGTTG GGATTCTCTG      691 OK
2468
2469  TGCAGAGACT TTGGTTCCCC ACGCAAGCCC TGGGGCTTGG AAGAAGCACA TGACCGTACT      751
2470
2471  CTGCGTGGGG CTCCACCTCA CACCCACCCC TGGGCATCTT AGGACTGGAG GGGCTCCTTG      811
2472
2473  GAAAACTGGA AGAAGTCTCA ACACTGTTTC TTTTTC      848
2474
2475

```

3000 (2) INFORMATION FOR SEQ ID NO:45:

```

3001
3002      (i) SEQUENCE CHARACTERISTICS:
--> 3003      (A) LENGTH: 4999 base pairs
3004      (B) TYPE: nucleic acid
3005      (C) STRANDEDNESS: single
3006      (D) TOPOLOGY: linear
3007
3008      (ii) MOLECULE TYPE: DNA

```

5000 shown (next page)

RAW SEQUENCE LISTING  
PATENT APPLICATION US/08/962,560DATE: 02/26/1999  
TIME: 14:56:40

INPUT SET: S30840.raw

3009  
3010  
3011  
3012 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:  
3013  
3014 CCCTCTGGGC AAGCCGCCCC CCCCCACCC ATCTACCACA CACACACACA CACACACACA 60  
3015  
3016 CACACATTCA GACCTTGGGG CAAAAACAAA GCAAATAAC AACAAACAAA AACTGCCTG 120  
3017  
3018 TGGAAAGTCC TTA CTTCAGG AAGGTTGGCA GATGAGGAGC AAGGGAACAT TTTATCAGGA 180  
3019  
3020 CTGCCACAAA GGAGTCTTTT TTTTAAATGG TTTTCAAGA CAGGGTTTCT CTGTATAGCC 240  
3021  
3022 CTGGCTGTCC TGGAGCTCAC TTTGTAGACC AGGCTGGCCT CGAACTCAGA AATTCGCCTG 300  
3023  
3024 CCTCTGCCTC CTGAGTGCTG GGATTAAAGG CGTGCAGCAC CATGTCCAAC TGGCATTTC 360  
3025  
3026 TCAATTAAGG TTCGTTCTT TCAGATAACT CTAGGTTCTG GGTCAAGCTG ACACAAGGCT 420  
3027  
3028 ACACAGCACA GTTTGTATGC CACATTCAGT TCAGAAGACA CCCAACCTCC CTGGAAGTGG 480  
3029  
3030 AACTTATGCA CATTGTGAG CTTCCACTTG GGAGTGGGAA CCTGAACTGG GTCCTCTGCA 540  
3031  
3032 AGAGCAGCCG TGCTCTTAAC TGCTGAGCCA TTTCAGCAGC CTCACATCAG AATTAAGTTA 600  
3033  
--> 3034 GAAATTAGCCG GGTATGAATC ATACCCTTAG AATCCTAGCA TCTGAAAGCA GAGCTAAGCA 660 661  
3035 AAACAGGGAT TCAAGACCAG CTCTTGGCTA CAGAGCCCGT CCTGTCCTAG GATGGGCTAC 720 721  
3036  
3037 AAGAGACTAT TTCAAAGCCA TCCAAACAAC AATAACTACA ACAACAACAA GGTTAAAATT 780  
3038  
3039 AGGCTGGGCA CAGGGTACAC ACCTTTAATG CCAACACTCA GGAGGCAGAG GCAGGCTGAT 840  
3040  
3041 CAGTGTGAGT TTGAGTTCAA CGTGGTCTAC ATAGGGAGTT CTAGGCCAGC AGAGGTTACA 900  
3042  
3043 GTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCACACA CACACACACA CACACACACA 960  
3044  
3045 CACACACACA CACACACGGT GGCATTATGG GATTTTTTTT GGATAAGGTT TCTCTGTCTA 1020  
3046  
3047 GCCCTGGCAT AGATTCACTC TGTAAGTAG GCTAGCCTTG AACTCAGAGA TCCGCCTGCC 1080  
3048  
3049 TCTGCCTCCC AAGTGCTGGG ATTATAGGTG TTGCACCACC ACTGCCAGC CACTTTGGGA 1140  
3050  
3051 TTTTGAACCT GTTATCAAGA GGCTTTCGAG GAGGTCAAAC TTCAACAGCA ACCTCTCCAT 1200  
3052  
3053 GATAATGTAG CTAATGATCA AACGACACTC AAAACTTAAC CCTTAAAGCA CACATCCACC 1260  
3054  
3055 AGACAGCGTG CCCACTCGTA GTTCCATTAC TCAGGAGGCT GAAGCAGGAG GATGAAGGAC 1320  
3056  
3057 TAAGGCTTCA GCAACCTAGG GAGCCGCAGG GGACAGTAGT CTCAATCCCT ACATTCTCCT 1380  
3058  
3059 GAACACAGGA GCAGGAGTTC AGGAAGGGTG TCAAGGCCGC TTACTGATCT TAGGGCCTCA 1440  
3060  
3061 GGAATGACTA GCTCAGGCAG AGAGAGCAAA GGTCTCCAGT GGAGAAGTCT ACACACACAC 1500

660 661  
720 721  
↓  
hunter  
off

3062							
3063	ACACACACAC	ACACACACAC	ACACACACAC	AGAATCCAAG	GCGATGACGT	CATCAAAGGG	1560
3064							
3065	TTAATTCTAG	TCTGGGATGG	GGGGGAGGGT	GGGGCACGCA	GCTGTCAGGT	GGCTTTGGAA	1620
3066							
3067	AAATAAACTG	CTGAAGAGTC	TGACGCCAGG	GAGTCCTGGG	AGGGACAAGA	GGTTACCCAC	1680
3068							
3069	TCAAAGAGTG	TGCTCCACAA	AGCATGCGCG	CTTGTCCACG	TCTGGAGTCG	TCACTTATTT	1740
3070							
3071	TTTGCCTGGA	TTCTTTGTAG	CCGGTGGGTT	CTCAAGGCGG	TAAGTGGTGT	GGCCGCCGTG	1800
3072							
3073	GTCTGGGAGG	TGACGATAGG	GTTAATCGTC	CACAGAGCCC	AGGGGCGGAG	CGCGGGCGGG	1860
3074							
3075	CGTCCGCAGC	CCCCTGGAG	CCGGAAGCAG	TGGCTGGTCA	GGGGCGCTTC	TAGCCTTCCC	1920
3076							
3077	TATCTGTACT	TCCACAGAGG	TCTCTGCGAG	CTAGGGGGAC	AGTGAGGTGC	GGGGTAGGGG	1980
3078							
3079	CCCGGCGTTA	GAGCCAGCAA	GGGGACGGTT	CACGGTAAGG	TCTGAGGGAG	AGAGAGCTCC	2040
3080							
3081	TGAGAAACTT	GGGGGGCGCG	ACACAGATAG	GGTGAAAGCA	GAGTGATAGA	CCTGGGATGG	2100
3082							
3083	TTAGGGGACC	AAGGGAAGAC	CAGGCTGGTT	GGCATAACC	GGTGAACGGA	TGGGAGTCCT	2160
3084							
3085	AGGGAAAGAT	GATGCGCCTA	ACAGTCCTTT	CTGTCTCCAC	ACCACTCCAG	GGGACGATCC	2220
3086							
3087	GGAGCTCAAC	TTTCAAAAGC	GAGACGCCCC	AGCAAGCCTG	TTTTGAGAAG	TTCTTCAGCG	2280
3088							
3089	GCTCTCCTCA	TGGGCCAGAC	GGCCCTGGCA	AGGGGCAGCA	GCAGCACCCC	TACCTCGCAG	2340
3090	GCTCTGTACT	CGGACTTCTC	TCCTCCCGAG	GGCTTGGAGG	AGCTCCTGTC	TGCTCCCCCT	2400
3091							
3092	CCTGACCTGG	TTGCCCAACG	GCACCACGGC	TGGAACCCCA	AGGATTGCTC	CGAGAACATC	2460
3093							
3094	GATGTCAAGG	AAGGGGTCT	GTGCTTTGAG	CGGCGCCCTG	TGGCCCAGAG	CACTGATGGA	2520
3095							
3096	GTCCGGGGGA	AACGGGGCTA	TTCGAGAGGT	CTGCACGCCT	GGGAGATCAG	CTGGCCCCTG	2580
3097							
3098	GAGCAAAGGG	GCACACACGC	CGTGGTGGGC	GTGGCCACCG	CCCTCGCCCC	GCTGCAGGCT	2640
3099							
3100	GACCACTATG	CGGCGCTTTT	GGGCAGCAAC	AGCGAGTCCT	GGGGCTGGGA	TATTGGGCGG	2700
3101							
3102	GGAAAATTGT	ATCATCAGAG	TAAGGGCCTC	GAGGCCCCCC	AGTATCCAGC	TGGACCTCAG	2760
3103							
3104	GGTGAGCAGC	TAGTGGTGCC	AGAGAGACTG	CTGGTGGTTC	TGGACATGGA	GGAGGGGACT	2820
3105							
3106	CTTGGCTACT	CTATTGGGGG	CACGTACCTG	GGACCAGCCT	TCCGTGGACT	GAAGGGGAGG	2880
3107							
3108	ACCCCTCTATC	CCTCTGTAAG	TGCTGTTTGG	GGCCAGTGCC	AGGTCCGCAT	CCGCTACATG	2940
3109							
3110	GGCGAAAGAA	GAGGTGAGAT	ACGGACTAGG	TGTGGGGAGA	TCACTACTCT	TGGCAATGGT	3000
3111							
3112	TTGGGCTGGA	AACTCATGGT	TGGAGCACAG	GAAGTAGGCT	TCTTGTCACT	TTGGCCTGTC	3060
3113							
3114	ACTTAGATGG	CCTTGGATCT	AGCTTCACTC	CCAATCCCTA	TGGGATGTGA	TGCACAAATT	3120

**INPUT SET: S30840.raw**

3115							
3116	CAGAGCCTTT	GGGTCTCCCT	CAGCTGAGGT	GGCGGTGGAA	ATGGAGGAAG	AAGGAAGGGT	3180
3117							
3118	GCCTGAGCAG	GATCTCAAGT	TCAAGGATGC	CTGGAGTTGC	TTACTTACCT	TGTCTTCCTT	3240
3119							
3120	CTCTCTCCGC	AGTGGAGGAA	CCACAATCCC	TTCTGCACCT	GAGCCGCCTG	TGTGTGCGCC	3300
3121							
3122	ATGCTCTGGG	GGACACCCGG	CTGGGTCAAA	TATCCACTCT	GCCTTTGCCC	CCTGCCATGA	3360
3123							
3124	AGCGCTATCT	GCTCTACAAA	TGACCCAGTA	GTACAGGGTG	TGCTGGCACC	CTACCGTGGG	3420
3125							
3126	GACAGGTGGA	GAGGCACCCG	CTGGCCTAGA	CAACTTTAAA	AAGCTGGTGA	AGCTGGGGGG	3480
3127							
3128	GGGGGGCTGG	ACCCCTTCAC	CTCCCCTTCT	CACAGGAGCA	AGACATATAG	AAATGATATT	3540
3129							
3130	AAACACCATG	GCAGCCTGGG	ACAAAGAGGT	TTTTGAAGTA	AAAAATGAGA	TGTATTGTCA	3600
3131							
3132	CAACCTGTTT	CATTATTGTT	TTTTGTTTTG	TTTTACACTC	CCCCACCCCA	GGCTAGAGCC	3660
3133							
3134	CCATCACTGT	CTTAAGGAAT	TATGACAACC	CACAAAGCTC	AGGCCCAGGT	GTTTATTTCC	3720
3135							
3136	CTTACATGTA	GGATGGTTCA	CAAACACAAT	ACAGGGGCTT	TGGCACCGTG	GGGGAGGGGA	3780
3137							
3138	CTATCCCAGG	CCTCTTAGGG	TCTCATGTAT	ACCGAATTCA	GACCCGAAAG	CTCTGAATTT	3840
3139							
3140	CTGCATCAGA	CATCCAGTAG	AACTTGGGAG	TGAAGCTAGA	GCCAAGGCCA	TCTAAGTGAC	3900
3141							
3142	AGGCCAAAGT	GACACGAAGC	CACTTCCCTG	TGCTCCAACC	ATGAGTTTCC	AGCCCCAAACC	3960
3143							
3144	AATGGAAGGT	GATTTCACTT	GTCAGGGCCC	AAAGGGACAG	TCAGTTCTAC	TCCCTCCCCCT	4020
3145	CACTAGGAGC	CACCTTGGTG	ACAGTTGATT	CTACCCACTG	TAAGTGGTAA	AGGGATTGGC	4080
3146							
3147	CTGGTCCCAA	CCATAATAGG	GCGGTGGAAA	CGGCTCAGGA	GGGTACAGCG	TGGATTAGGC	4140
3148							
3149	CACAAGATGG	GGCAGATGAT	GTCATCAGAA	GCATGTGACC	GGTGGGAGCA	GTTACTAAAC	4200
3150							
3151	TTCTGGGCAA	CCTAGTCCAT	GCTATGCAGG	CAGGTAGAGG	GATGGGCAGT	GCTCATTGTT	4260
3152							
3153	TGGCATTGAT	GATGTCCACA	AATTCAAGCT	TGAGAGATGC	GCCACCCACA	AGGAAGCCGT	4320
3154							
3155	CCACGTCAGG	CTGGCTTGCC	AGCTCTTTGC	AGGTTGCTCC	AGTCACAGAA	CCTGTACCAG	4380
3156							
3157	GAACAAGAAG	ACAGTTTGGT	CAGGTCTATG	ATCAGAACAC	TTAAGCCCCA	CCTCTCTGTG	4440
3158							
3159	CAAGGCAGCC	TCAGTCTGTC	TTAGCCCATT	TCCGTCTTAG	CTAGAGCCAA	AGCCACTCAC	4500
3160							
3161	CTCCATAAAT	GATCCGGGTG	CTCTGAGCCA	CCCCATCATT	GACATTGGAT	TTCAGCCATC	4560
3162							
3163	CCCGGAGCTT	CTCGTGTACT	TCCTGTGCCT	AGAAGGAGGA	GGCAGAGCTA	CTAAGTAAGC	4620
3164							
3165	TCCTTCCTAT	CTATCATTCA	AGGAGTAAAA	ACCACTGGTT	CTCACATAGA	GTTGAGTTTC	4680
3166							
3167	CAGAAAAGCC	CCGGGACCAG	AGAGTGGCAA	GGCTCCAATC	CCACCAGGCT	TGGAATGAAC	4740

number  
Mr

**INPUT SET: S30840.raw**

[illegible]

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 421 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Ala Ser Phe Pro Pro Arg Val Asn Glu Lys Glu Ile Val Arg Ser  
1 5 10 15  
Arg Thr Ile Gly Glu Leu Leu Ala Pro Ala Ala Pro Phe Asp Lys Lys  
20 25 30  
Cys Gly Gly Glu Asn Trp Thr Val Ala Phe Ala Pro Asp Gly Ser Tyr  
35 40 45  
Phe Ala Trp Ser Gln Gly Tyr Arg Ile Val Lys Leu Val Pro Trp Ser  
50 55 60  
Gln Cys Arg Lys Asn Phe Leu Leu His Gly Ser Lys Asn Val Thr Asn  
65 70 75 80  
Ser Ser Cys Leu Lys Leu Ala Arg Gln Asn Ser Asn Gly Gly Gln Lys  
85 90 95  
Asn Lys Pro Pro Glu His Val Ile Asp Cys Gly Asp Ile Val Trp Ser  
100 105 110  
Leu Ala Phe Gly Ser Ser Val Pro Glu Lys Gln Ser Arg Cys Val Asn  
115 120 125  
Ile Glu Trp His Arg Phe Arg Phe Gly Gln Asp Gln Leu Leu Leu Ala  
130 135 140  
Thr Gly Leu Asn Asn Gly Arg Ile Lys Ile Trp Asp Val Tyr Thr Gly  
145 150 155 160  
Lys Leu Leu Leu Asn Leu Val Asp His Ile Glu Met Val Arg Asp Leu  
165 170 175  
Thr Phe Ala Pro Asp Gly Ser Leu Leu Leu Val Ser Ala Ser Arg Asp  
180 185 190  
Lys Thr Leu Arg Val Trp Asp Leu Lys Asp Asp Gly Asn Met Val Lys  
195 200 205  
Val Leu Arg Ala His Gln Asn Trp Val Tyr Ser Cys Ala Phe Ser Pro  
210 215 220  
Asp Cys Ser Met Leu Cys Ser Val Gly Ala Ser Lys Ala Val Phe Leu  
225 230 235 240  
Trp Asn Met Asp Lys Tyr Thr Met Ile Arg Lys Leu Glu Gly His His  
245 250 255  
His Asp Val Val Ala Cys Asp Phe Ser Pro Asp Gly Ala Leu Leu Ala  
260 265 270

08/9/2560

*see next page*



08/962, 560

Thr Ala Ser Tyr Asp Thr Arg Val Tyr Val Trp Asp Pro His Asn Gly  
275 280 285

Asp Leu Leu Met Glu Phe Gly His Leu Phe Pro Ser Pro Thr Pro Ile  
290 295 300

Phe Ala Gly Gly Ala Asn Asp Arg Trp Val Arg Ala Val Ser Phe Ser  
305 310 315 320

His Asp Gly Leu His Val Ala Ser Leu Ala Asp Asp Lys Met Val Arg  
325 330 335

Phe Trp Arg Ile Asp Glu Asp Cys Pro Val Gln Val Ala Pro Leu Ser  
340 345 350

Asn Gly Leu Cys Cys Ala Phe Ser Thr Asp Gly Ser Val Leu Ala Ala  
355 360 365

Gly Thr His Asp Gly Ser Val Tyr Phe Trp Ala Thr Pro Arg Gln Val  
370 375 380

Pro Ser Leu Gln His Ile Cys Arg Met Ser Ile Arg Arg Val Met Ser  
385 390 395 400

Thr Gln Glu Val Gln Lys Leu Pro Val Pro Ser Lys Ile Leu Ala Phe  
405 410 415

Leu Ser Tyr Arg Gly  
420

\* Delete  
this - otherwise,  
processed as  
an error

SEQUENCE VERIFICATION REPORT  
PATENT APPLICATION US/08/962,560

DATE: 02/26/1999

TIME: 14:56:42

INPUT SET: S30840.raw

Line	Error	Original Text
5	Mandatory Value Not Present	(i) APPLICANT:
34	Wrong application Serial Number	(A) APPLICATION NUMBER: US APPLICATION NO. 0
64	Entered (20) and Calc. Seq. Length (0) differ	(A) LENGTH: 20 base pairs
80	Entered (20) and Calc. Seq. Length (0) differ	(A) LENGTH: 20 base pairs
95	Entered (1236) and Calc. Seq. Length (1235) differ	(A) LENGTH: 1236 base pairs
110	# of Sequences for line conflicts w/ running total	CGAGGCTCAA GCTCCGGGCG GATTCTGCGT GCCGC
111	# of Sequences for line conflicts w/ running total	GGCCTGTGCC ACCCGGACGC CCGGCTCACT GCCTC
113	# of Sequences for line conflicts w/ running total	GACGCTATGG CCCACCCCTC CAGCTGGCCC CTCGA
116	# of Sequences for line conflicts w/ running total	ATG GTA GCA CGC AAC CAG GTG GCA GCC GAC A
575	Wrong Or Missing Strandedness Value	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:
661	Wrong Or Missing Strandedness Value	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:
907	Entered (421) and Calc. Seq. Length (422) differ	(A) LENGTH: 421 amino acids
995	Stop Codon at end of sequence removed - no error	(2) INFORMATION FOR SEQ ID NO:15:
2400	Entered (848) and Calc. Seq. Length (793) differ	(A) LENGTH: 848 base pairs
2467	# of Sequences for line conflicts w/ running total	AATTAAGCTA CCATGAAAAG AAGAGGAAAA GTGA
3003	Entered (4999) and Calc. Seq. Length (5000) differ	(A) LENGTH: 4999 base pairs
3034	# of Sequences for line conflicts w/ running total	GAAATTAGCCG GGTATGAATC ATACCCTTAG AATC